

Achieving clinical pictures of melanocytic skin neoplasms

before the histopathologic examination

does not introduce a bias in the diagnostic process

Gerardo Ferrara,^a MD, Giorgio Annessi,^b MD, Zsolt Argyenyi,^c MD, Giuseppe Argenziano,^d MD, Helmuth Beltraminelli,^e MD Rino Cerio,^f MD, Lorenzo Cerroni,^g MD, Carlo Cota,^h MD, Stefano Simonetti,ⁱ MD, Catherine M. Stefanato,^j MD FRCPATH, Iris Zalaudek,^g MD, Harald Kittler,^k MD, H. Peter Soyer,^l MD

Conflicts of interest: none to be declared

Role of funding source: none

Word count: 1807

Figures: 2

Tables: 1

References: 28

Running title: Clinicopathologic correlation

Keywords: Melanocytic skin neoplasms - Histopathologic diagnosis - Clinicopathologic correlation - Clinical information - Dermoscopy

^aAnatomic Pathology Unit, Gaetano Rummo General Hospital, Benevento, Italy

^bDermatopathology Unit, Istituto Dermopatico dell'Immacolata, Rome, Italy

^cDepartment of Dermatology, University of Washington, Seattle, Seattle, U.S.A. (Prof Z Argyenyi, MD),

^dDepartment of Dermatology, Arcispedale 'S. Maria Nuova' Reggio Emilia, I

^eDepartment of Dermatology, University of Bern, CH

^fDepartment of Dermatology, University of London, London, U.K. (Prof R Cerio, MD),

^gDepartment of Dermatology, Medical University of Graz, Graz, Austria

^hDermatopathology Unit, San Galligano Institute, Rome, Italy

ⁱDepartment of Dermatology, University of Perugia, Perugia, Italy

^jDepartment of Dermatopathology, St John's Institute of Dermatology, St Thomas' Hospital, London, U.K.

^kDepartment of Dermatology, Division of General Dermatology, Medical University of Vienna, Austria

^lDepartment of Dermatology Group, School of Medicine, The University of Queensland, Brisbane, Australia

Correspondence to: Gerardo Ferrara – Department of Oncology - Anatomic Pathology Unit – 'Gaetano Rummo' General Hospital – Via dell'Angelo 1 – I82100 Benevento – ITALY

Phone +39082457315

Fax +39082457334

e-mail: gerardo.ferrara@libero.it

38

39 **Abstract**

40 **Background** – The *a priori* knowledge of the clinical picture of melanocytic skin
41 neoplasms (MSN) introduces a potential bias in the histopathologic examination.

42 **Methods** – Histologic slides from 99 MSN were circulated among ten histopathologists
43 with clinical expertise: five histopathologists had clinical images available after a ‘blind’
44 examination (Group 1); the other five had clinical images available before microscopic
45 examination (Group 2). Data from the two groups were compared regarding consensus
46 diagnosis (CD: a diagnosis in agreement by ≥ 4 histopathologists/group), chance-corrected
47 interobserver agreement (Fleiss’ k), and level of diagnostic confidence (LDC: a 1 to 5
48 arbitrary scale indicating ‘increasing reliability’ of any given diagnosis).

49 **Results** – Compared with Group 1 histopathologists, Group 2 achieved a lower number of
50 CD (84 vs. 90) but a higher k value (0.74 vs. 0.69) and a greater mean LDC value (4.57
51 vs 4.32). The same CD was rendered by the two groups in 81/99 cases. Spitzoid MSN
52 were most frequently controversial for both groups.

53 **Conclusions** – The histopathologic interpretation of MSN seems to be not influenced by
54 the knowledge of the clinical picture before histopathologic examination..

55

56 Introduction

57 Whether any diagnostic test should be read together with clinical information has
58 been debated since 1963, when Schreiber¹ suggested the clinical information as a way to
59 improve the accuracy of chest X-ray evaluation. Nonetheless, the impact of the clinical
60 information in diagnostic cytohistopathology has been addressed by very few studies.²⁻⁵ A
61 commonly used argument against the 'interdisciplinary' approach is that clinical
62 information may bias the reading:⁶⁻⁷ in order to minimize such a bias, the microscopic
63 examination should be carried out first blind to the clinical information and then in light of
64 them:^{4,5,8} thus, *perception* (identification of abnormal areas and their features) would result
65 unbiased by the clinical information, whereas the latter should help the final *interpretation*
66 (attribution of the abnormalities to an entity). By following such a methodology, two recent
67 studies performed in the field of dermatopathology have shown that the dermatopathologic
68 diagnosis starts as a perception of microscopic criteria which can work as such, but is
69 finally a clinically-aided interpretation.⁴⁻⁵ A major problem, however, is the potential bias
70 born by an *a priori* knowledge of the clinical picture, as happens when clinical
71 dermatologists evaluate the histologic specimens from their own patients.^{7,9-11}

72 In a previous study,⁴ the histologic slides from 99 clinically atypical melanocytic skin
73 neoplasm (MSN) were submitted to ten histopathologists, five of these with clinical
74 expertise ('clinical histopathologists'); in order to evaluate the diagnostic impact of the
75 single clinical data, clinical information were given with a five-step procedure (no
76 information; age/sex/location; clinical diagnosis; clinical image; dermoscopic image). Steps
77 1-3 of such a procedure excluded the bias of the knowledge of the clinical picture before
78 histopathology. We now introduced such a bias by submitting the same dataset from the
79 previous study⁴ to five clinical histopathologists who were requested to look at the
80 histologic slides only after having all pertinent clinical information. The new data were
81 compared with those given by the former five clinical histopathologists.

Methods

After an *ad hoc* authorization obtained by the patients or their guardians, 99 consecutive cases of clinically/dermoscopically atypical MSN were submitted to two groups of clinical histopathologists:

Group 1 (ZA, RC, LC, HK, HPS): histopathologists having a stepwise access to clinical information in the course of their microscopic evaluation;

Group 2 (GAn, HB, CC, SS, CMS): histopathologists having full knowledge of the clinical picture before the microscopic examination.

Each panelist is almost equally involved in routine clinical and histopathologic work. Only two of the above panelists (LC and HPS) had worked for several years at the same Institution.

For each case, a single hematoxylin-eosin stained slide, accurately checked for its technical and diagnostic adequacy, was provided to each panelist.

All the clinical information concerning the selected cases was included into a FileMaker Pro 7TM (FileMaker Inc.)-generated database. For each case, Group 1 histopathologists were requested to evaluate the microscopic slide according to a five-step procedure: i) no information; ii) knowledge of age and gender of the patient and location of the lesion; iii) clinical diagnosis; iv) clinical image; v) dermoscopic image. Group 2 histopathologists had all information available before microscopic examination.

The influence derived on the diagnosis of MSN from the knowledge of the full clinical information before the histopathologic examination was checked by comparing the data provided by Group 2 histopathologists with the data provided by Group 1 histopathologists. The parameters evaluated were: the consensus diagnosis (CD), the chance-corrected interobserver agreement, and the mean level of diagnostic confidence (LDC). For statistical analysis, all the diagnoses were grouped into two ratings: 'melanoma' and 'nevus'.

CD was defined as a diagnosis made in agreement by at least four out of five panelists per group. This have been recently proposed as a 'surrogate' gold standard when follow-up data can give little information (as in this case: see the Results section).^{12,13} Unanimous diagnoses were a subgroup of CDs; the comparison between the number of both unanimous diagnoses and CDs given by Group 1 and Group 2 histopathologists was performed with McNemar's test.¹⁴ The given *p*-value is one-tailed and a *p*-value of <0.05 indicates statistical significance.

Case by case, the histopathologists were asked to log into the database also a *level of diagnostic confidence* (LDC), namely, the probability, as scored according to an arbitrary scale, that they 'subjectively' attributed to the given diagnosis.¹⁵ The LDC scale was structured into five levels:

LDC 1 – No diagnostic certainty: no diagnosis can be made.

LDC 2 – Low diagnostic certainty: a diagnosis is felt as slightly more likely.

LDC 3 – Moderate diagnostic certainty: a diagnosis is favoured, but with some elements of doubt.

LDC 4 – High diagnostic certainty: a diagnosis is strongly favoured.

LDC 5 – Absolute diagnostic certainty: no other diagnosis is possible.

The interobserver agreement among the observers was calculated using the *k* statistics for multiple ratings introduced by Fleiss.¹⁶⁻¹⁸ Given *n*=5 as the number of panelists, *k* values range between +1 (perfect agreement) and <0 (perfect disagreement); values greater than 0.75 represent an excellent agreement; values lower than 0.40 a poor agreement. For the comparison of LDC values a Normal z-test for dependent samples was used.¹⁹ The given *p*-value is 2-tailed and a *p*-value <0.05 indicates statistical significance.

132

133 **Results**

134 The study included 99 cases from 96 patients (M:F=0.6:1; age range: 10-78 years; mean
 135 age: 43.3 years; median age: 42 years). The most common location was the back (40
 136 cases), followed by the lower limbs (16 cases) and by the upper limbs (12 cases). The
 137 referring histopathologist had diagnosed 54 cases as 'nevus' and 45 cases as 'melanoma'.
 138 Follow-up (range: 24-96 months; mean: 54.5 months) was available in 65 cases, 30 of
 139 which originally diagnosed as melanoma. Ten of these cases underwent sentinel node
 140 biopsy, which was negative in all instances; an adverse clinical outcome was recorded in
 141 five cases (four sentinel node negative cases in which distant metastases developed; one
 142 0.80 mm-thick melanoma in which a nodal metastasis detected 40 months after excision).
 143 All the metastasizing cases received a CD of melanoma from both groups of observers.

144 Within Group 1 histopathologists, 65 cases had a unanimous diagnosis and 25
 145 cases had only one discrepant diagnosis, with a CD obtained in 90/99 cases. The k value
 146 was 0.69 and the LDC 4.32 ± 0.59 .

147 Group 2 histopathologists diagnosed 75/99 MSN in unanimity and 9 MSN with only
 148 one discrepant diagnosis: therefore, a CD was achieved in 84/99 cases. The k value was
 149 0.74; the LDC was 4.57 ± 0.39 . Thus, compared with data from Group 1, there was a lower
 150 number of CD, but a greater chance-corrected interobserver agreement and a greater
 151 mean LDC value. Remarkably, the differences among the number of unanimous
 152 diagnoses as well as the number of CDs generated by the two groups was not statistically
 153 significant (McNemar's test: $p=0.078$ for the differences in the number of unanimous
 154 diagnoses; $p=0.18$ for the differences in the number of CDs), whereas the difference
 155 between the mean LDC values was highly significant (Normal z-test $p<0.001$).

156 By comparing case by case the CD generated by the two groups of observers, 81
 157 cases resulted with the same CD; five cases were controversial for both groups; six cases

158 were controversial only for Group 1; and seven cases were controversial only for Group 2.
159 The five most controversial cases showed Spitzoid (3/5; Fig. 1) or lentiginous (2/5; Fig. 2)
160 features. Overall, the most common morphologic patterns of controversial cases (Table 1)
161 were Spitzoid (8/18 cases), lentiginous (3/18), and regression(-like). Among these
162 controversial cases, there was no association between a given morphologic pattern and
163 the protocol of microscopic observation followed.

164

165

166

167 Discussion

168 This study demonstrates that the histopathologist who knows the clinical picture of MSN
169 before the histopathologic examination is not biased in the final interpretation and is
170 actually strengthened in the diagnostic process.

171 It has been demonstrated that full clinical information provided after a first
172 microscopic examination accomplished blind to the clinical data can aid the final
173 interpretation for both MSN⁴ and inflammatory dermatoses.⁵ It is still disputable, however,
174 whether a bias is introduced in the diagnostic process when the full knowledge of the
175 clinical picture is achieved before histopathologic examination. This issue also involves the
176 neverending debate regarding the histopathologic practice by dermatologists and the level
177 of training required them to sign out dermatopathology cases.^{7,20}

178 We have compared some parameters (CDs, chance-corrected interobserver
179 agreement, mean LDC) provided in a series of 99 MSN by two groups of histopathologists
180 with clinical expertise: one of these (Group 1) had clinico-dermoscopic images available
181 before the microscopic examination, whereas the other (Group 2) had all clinical
182 information available *a priori*. Within the Group 2 there was a lower number of CDs: this
183 finding, although not statistically significant, could suggest that a *perception* influenced by
184 the knowledge of the clinical picture does necessarily translate into a more homogeneous
185 *interpretation*. Therefore, the extreme scenario of a clinical prejudice forcing the
186 histopathologic diagnosis²¹ is unlikely. Instead, the *k* value generated by Group 2
187 histopathologists was greater than Group 1: this means that the influence of the clinical
188 pictures is similar, regardless the moment in which they are introduced into the diagnostic
189 process.

190 A further interesting finding is the greater LDC found within Group 2
191 histopathologists, a finding which was statistically significant. This could imply that the
192 clinical picture, albeit being unable to bias the final *interpretation*, can aid the *perception*

and give greater strength to the diagnostic process. Parenthetically, among controversial cases, there was no association between a given morphologic pattern and the protocol of microscopic observation followed. As expected, controversies in Spitzoid MSN^{4,22} (an example is given in Fig. 1) were found to be sizable, independent from the moment in which the histopathologist is aware of the clinical picture. Along with the results of our previous study,⁴ we also found some diagnostic controversy regarding lentiginous MSN (an example is given in Fig. 2): all the disputable lesions of this category were removed from the back of middle-aged patients (data not shown). It has been previously demonstrated that the differential diagnosis between lentiginous 'dysplastic' nevus and lentiginous melanoma can be significantly aided by the clinico-dermoscopic digital monitoring of the lesions, since lentiginous melanoma continuously, but slowly grows and remains 'in situ' over years or even decades.²³⁻²⁵ In the present study, however, data regarding the 'E' (=evolution) criterion of the ABCDE clinical alphabet of melanoma²⁶ were not provided, and this must be underlined as a limitation in the clinicopathologic evaluation made by the panelists. Lesions with regression(-like) features^{24,27,28} have been found somewhat controversial, but to a lesser extent than expected. This can be due to the fact that all the MSN included in the present study had undergone macroscopic sampling according to their dermoscopic features, a procedure which can help the diagnostic evaluation of these lesions by highlighting their most atypical features.²⁴

In conclusion, there is evidence that clinical pictures of MSN can aid their *interpretation* and that the latter is not biased if the clinical pictures are available before the microscopic examination. Good lines of communication between dermatologists and histopathologists are always desirable; when the dermatologist and the histopathologist are the same person, the communication is best and, if intellectually honest, is probably not misleading.

References

- 1.Schreiber MH. The clinical history as a factor in roentgenogram interpretation. JAMA 1983;185:399.
- 2.Abbey LM, Kaugars GE, Gunsolley JC, et al. The effect of the clinical information on the histopathologic diagnosis of oral epithelial dysplasia. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1998;85:74.
- 3.Raab SS, Oweity T, Hughes JH, et al. Effect of clinical history on diagnostic accuracy in the cytologic interpretation of bronchial brush specimens. Am J Clin Pathol 2000;114:78-83.
- 4.Ferrara G, Argenyi Z, Argenziano G, et al. The influence of the clinical information in the histopathologic diagnosis of melanocytic skin neoplasms. PLoS ONE 2009; 4: e5375. doi:10.1371/journal.pone.0005375. Epub 2009 Apr 30.
- 5.Cerroni L, Argenyi Z, Cerio R, et al. Influence of evaluation of clinical pictures on the histopathologic diagnosis of inflammatory skin disorders. J Am Acad Dermatol 2010; 63:647.
- 6.Shitabata PK. Do clinical histories submitted on biopsy requisition sheets bias the Pathologists? <http://www.dermpathmd.com>
- 7.Ackerman AB (2005) Dermatologist not equal dermatopathologist: no place in a profession for pretenders. J Am Acad Dermatol 2005; 53:698.
- 8.Griscom NT (2002) A suggestion: look at the images first, before you read the history. Radiology 223:9-10.
- 9.Moy R, The reason that dermatologists should not send all their slides to dermatopathologists is a scope of practice argument, not an ethical argument. J Am Acad Dermatol 2005, 53:700.
10. Glogau RG, Collegiality and dermatology. J Am Acad Dermatol 2005, 53:701.

- 245 11. Ferrara G, Crisman G. F1000 Prime Recommendation of [Cerroni L, Argenyi
246 Z, Cerio R, et al. Influence of evaluation of clinical pictures on the histopathologic
247 diagnosis of inflammatory skin disorders. J Am Acad Dermatol 2010; 63:647].
248 <http://www.f1000.com/prime/5557956#eval5525054>
- 249 12. Coggon D, Martyn C, Palmer KT, Evanoff B. Assessing case definitions in
250 the absence of a diagnostic gold standard. Int J Epidemiol 2005;34:949.
- 251 13. Ferrara G, Tomasini C, Argenziano G, Zalaudek I, Stefanato CM. Small-
252 diameter melanoma: toward a conceptual and practical reappraisal. J Cutan Pathol 2012;
253 39:721.
- 254 14. McNemar Q. Note on the sampling error of the difference between correlated
255 proportions or percentages". *Psychometrika* 1947;12: 153.
- 256 15. Meghini C, Sebastiani F, Straccia U, Thanos C. A model of information
257 retrieval based on a terminological logic. In: Proceedings of SIGIR-93, 16th ACM
258 International Conference on Research and Development in Information Retrieval.
259 Pittsburgh, PA, 1993;298.
- 260 16. Fleiss J L, Cohen J, Everitt BS. Large sample standard errors of kappa and
261 weighted kappa. Psychol Bull 1969; 72:323.
- 262 17. Fleiss JL. Measuring nominal scale agreement among many raters. Psychol
263 Bull 1971; 76:378.
- 264 18. Fleiss JL. Statistical Methods for Rates and Proportions . 2nd ed. New York,
265 Wiley, 1981:2.
- 266 19. D'Ambra L. Verifica delle ipotesi. In: Lezioni di Inferenza Statistica. Rocco
267 Curto Editore, Naples, 2007;232.
- 268 20. Glusac EJ. Under the microscope: doctors, lawyers, and melanocytic
269 neoplasms. J Cutan Pathol 2003;30:287.
- 270 21. Grant-Kels JM. The whys and the wherefores of who reads

271 dermatopathologic slides. J Am Acad Dermatol 2005; 53:703.

272 22. Barnhill R The Spitzoid lesion: the importance of atypical variants and risk
273 assessment. Am J Dermatopathol 2006; 28:75.

274 23. Ferrara G, Zalaudek I, Argenziano G. Lentiginous melanoma: a distinctive
275 clinicopathological entity. Histopathology 2008; 52: 523.

276 24. Ferrara G, Argenziano G, Giorgio CM, Zalaudek I, Kittler H. Dermoscopic-
277 pathologic correlation: apropos of six equivocal cases. Semin Cutan Med Surg 2009; 28:
278 157.

279 25. Terushkin V, Dusza SW, Scope A, et al. Changes observed in slow growing
280 melanomas during long-term dermoscopic monitoring. Br J Dermatol 2012; 166:1213.

281 26. Abbasi NR, Shaw HM, Rigel DS. Early diagnosis of cutaneous melanoma.
282 Revisiting the ABCD criteria. JAMA 2004; 292: 2771.

283 27. Ferrara G, Argenziano G, Soyer HP, et al. Dermoscopic and histopathologic
284 diagnosis of equivocal melanocytic skin lesions. An interdisciplinary study on 107 cases.
285 Cancer 2002;95:1094.

286 28. Zalaudek I, Argenziano G, Ferrara G, et al. Clinically equivocal melanocytic
287 skin lesions withy features of regression: a dermoscopic-pathological study. Br J Dermatol
288 2004; 150:64.

289

290

291 Table 1 – Morphologic features of the most controversial cases of the present study

292 distributed according to the group

293

Panelists	Main morphologic pattern						
	<i>Spitzoid</i>	<i>Lentiginous</i>	<i>Regression(-like)</i>	<i>Nested</i>	<i>Halo</i>	<i>Congenital nevus-like</i>	Total
All	3	2	0	0	0	0	5
Group 1	2	0	2	1	1	0	6
Group 2	3	1	1	1	0	1	7
Total	8	3	3	2	1	1	18

294

Figure legends

Figure 1: A melanocytic lesion of the thigh in a 58-year-old woman. Dermoscopy is characterized by a striking asymmetry, with a reticular depigmentation associated with a dotted vascular pattern in its upper portion (A). Histopathologically, the neoplasm is compound, with a moderate epidermal hyperplasia (B), lack of circumscription (C), and with a Spitzoid cytomorphology (D). The differential diagnosis is between a Spitz nevus and an early invasive Spitzoid melanoma.

Figure 2: A flat melanocytic lesion of the back in a 58-year-old man. Dermoscopically there is an atypical pigment network with features of regression (A); histopathologically, the lesion is large, with partially preserved retiform epidermal hyperplasia (B), a prevailing single cell proliferation at the junction, involvement of the adnexal epithelium (C) and cytologic features of 'severe dysplasia' (D). The differential diagnosis is between a 'severely dysplastic' (Clark) nevus and a lentiginous melanoma in situ.²³